

UNITED STATES PATENT APPLICATION

of

Marko Kocic

Lidwine Martinot

Zoran Zvonar

for

**IMPROVED METHOD FOR JOINT DC OFFSET CORRECTION AND
CHANNEL COEFFICIENT ESTIMATION IN A RECEIVER**

IMPROVED METHOD FOR JOINT DC OFFSET CORRECTION AND CHANNEL COEFFICIENT ESTIMATION IN A RECEIVER

BACKGROUND OF THE INVENTION

5 The invention relates to the field of data receivers, and, in particular, to systems and methods for reducing DC offsets in wireless receivers.

 The basic challenge in detection of transmitted data comes from distortion caused by the channel. Typically, channel refers to a combination of all non-ideal effects that adversely affect the ability to correctly detect the transmitted data sequence at the receiver. Such effects
10 include, but are not limited to, effects of propagation (e.g. cable, radio channel etc.), non-linearities, or effects of transmit (TX) and receive (RX) filters. Current state-of-the-art detection techniques rely on the knowledge of the channel, which, therefore, must be estimated.

 The background of the invention will be described in the wireless communication
15 scenario, although invention can be used in any communication system. As is known in the art, in wireless communications received radio frequency signals are converted to baseband using various receivers. With a homodyne receiver, the received radio frequency signal is mixed with the local oscillator whose frequency is equal to carrier frequency of the received radio frequency signal to translate the carrier frequency to DC and thereby provide "direct
20 conversion" of the modulation on the received radio frequency signal to a modulation at DC. Hence, a homodyne receiver is sometimes referred to as a direct conversion receiver.

While such direct conversion receivers offer the advantage of flexible architecture and reduced cost, DC offset associated with such receivers is affecting the receiver performance. The level of the DC offset in direct conversion receiver may range from being significantly larger than the desired signal, i.e., modulation, to being comparable in level with desired signal. Thus, DC offset compensation techniques are typically required. To insure flexibility for different operating conditions, DC offset compensation can be part of the digital baseband portion of the receiver, preferably a part of the digital signal processing (DSP) firmware. One application of direct conversion receivers is in mobile or cellular communication systems. In such systems, the radio channels received signals also suffer from intersymbol interference (ISI) caused by multipath radio propagation and transmitter and/or receiver filtering. We further explain the underlying principle using the specific format of Time Division Multiple Access (TDMA) system as represented by GSM/GPRS/EDGE cellular standard.

The original GSM system, of which GPRS and EDGE are extensions, is a TDMA system, which means that a particular channel is active only during specified time periods. In this case, each channel is divided into eight time slots, with user being active at only every eight slots for voice transmission, or utilizing multiple slots for data transmission per standard specification.

GSM (including GPRS and EDGE) uses the midamble training sequence (a training sequence placed in the middle of the burst of data that occurs in a given slot) in every burst. The system parameters are chosen so that the change in Channel Impulse Response (CIR) is

fairly small during one burst so that adaptive CIR estimation is not needed. The midamble training sequence minimizes the distance (in time) between the known portion of the burst (the training sequence) and the ends of the burst, thereby also minimizing CIR change with respect to the CIR estimate obtained from the training sequence.

5 In EDGE system, the combined channel response usually consists of the linear TX Gaussian shaping filter (C0), actual physical channel, and TX/RX filters (analog and/or digital filters). A set of channel models is used for performance evaluation and covers a wide range of typical physical propagation channels that are likely to be encountered in practice as specified by the standardization bodies (ETSI, 3GPP). These channels are denoted by:

- 10 - static (no ISI or fading)
- TU3 (typical urban, velocity of the mobile station (MS) 3km/h, moderate multipath, low fading)
- TU50 (typical urban, velocity of the MS 50km/h, moderate multipath and fading)
- RA250 (rural area, MS velocity 250km/h, no multipath, severe fading) HT100 (hilly
- 15 terrain, MS velocity 100km/h, severe multipath, significant fading).

GSM training sequences (shared by GPRS and EDGE) have good autocorrelation properties. GSM/GPRS systems use GMSK modulation format, while EDGE introduces the new 8-PSK modulation format. In EDGE, training sequences for GMSK and 8-PSK modulation are identical, binary, and differ only due to different additional rotation applied to the signal in

20 two modulation schemes. GMSK effectively rotates a signal 90 degrees every symbol, while EDGE 8-PSK uses the rotation of $3\pi/8$ for every symbol. Properties of the training sequences

are used to simplify CIR estimation. For example, due to good autocorrelation properties of training sequences, simple cross-correlation can be used for CIR estimation.

In a typical data receiver realization in GSM/GPRS/EDGE system, the whole received burst is stored and the all-digital techniques may be applied to signal processing, including DC offset compensation. Thus, referring to FIG. 1, the data receiver stores the burst of data, $r(k)$, where $k=1 \dots N$ and N is the number of samples in the burst. Each burst includes a mid-amble having a known sequence of bits disposed between data, (i.e., information bits) as shown. Such known sequence of bits is used to aid in equalization and more particularly for enabling channel estimation, which includes timing information, length of the CIR, calculation of CIR taps as well as the alignment of the CIR estimate with respect to the largest magnitude CIR tap, and DC offset estimation. As shown in FIG. 1, an estimate of the DC offset, \hat{A} , is calculated. The estimated DC offset, \hat{A} , is subtracted from the received burst. The result, $r(k) - \hat{A}$, where $k=1 \dots N$, is processed to find an estimate of the h . The CIR estimate, \hat{h} , can be obtained by a variety of methods including cross-correlating $[r(k) - \hat{A}]$ with the known mid-amble bit sequence, using Least Square (LS) estimation as well as other advanced techniques.

The simplest method for channel estimation in GSM is cross-correlation. In cross-correlation CIR estimation, the inner 16 symbols of the training sequence are correlated with the received symbol. The received symbol at time k is given by:

$$y_n = \sum_{i=0}^{L-1} h_i t_{n-i} + z_n$$

where y_k is the received signal, h_k are CIR taps, t_k is the known training sequence (or data), while z_k is AWGN with variance σ^2 . The following notation is used to differentiate between the (known) training sequence and (generally unknown) data. Training sequence symbols are denoted by t_0 to t_{25} . Data for the left data burst are denoted d_0 to d_{57} , while those for the right data burst are denoted d_{58} to d_{115} . By cross-correlation with the inner 16 training sequence symbols:

$$\hat{h}_k = \frac{1}{16} \sum_{j=5}^{20} y_{k+j} t_j = \frac{1}{16} \sum_{j=5}^{20} t_j \left(\sum_{i=0}^{L-1} h_i t_{j-i+k} + z_{j+k} \right) = \frac{1}{16} \sum_{i=0}^{L-1} h_i \sum_{j=5}^{20} t_j t_{j-i+k} + \frac{1}{16} \sum_{j=5}^{20} t_j z_{j+k}.$$

Due to autocorrelation properties of the training sequences, $\sum_{j=5}^{20} t_j t_{j-i+k} = 0$, unless $i=k$ (provided $|k-i| < 5$),

$$\hat{h}_k = \frac{1}{16} \sum_{j=5}^{20} y_{k+j} t_j = h_k + \frac{1}{16} \sum_{j=5}^{20} t_j z_{j+k} = h_k + \Delta h_k.$$

Δh_k denotes CIR tap estimation error due to noise. The main limitation of cross-correlation CIR estimation is that it can only provide accurate estimates of up to 6 CIR taps. If CIR is longer than 6, and HT100 channel profiles span seven symbols when convolved with Gaussian pulse C0 (3 symbols span), cross-correlation CIR estimation produces inaccurate estimates for all but six initial taps.

For example, for CIR length of 6, in matrix notation, \hat{h} is given by:

$$\hat{h} = C(Th + z)$$

where

$$C = \begin{bmatrix} t_5 & t_6 & \cdots & t_{20} & 0 & \cdots & \cdots & 0 \\ 0 & t_5 & \cdots & \cdots & t_{20} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & & & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & t_5 & \cdots & \cdots & t_{20} \end{bmatrix}$$

5

$$T = \begin{bmatrix} t_5 & t_4 & \cdots & t_0 \\ t_6 & t_5 & \cdots & t_1 \\ \vdots & & & \vdots \\ t_{25} & \cdots & \cdots & t_{20} \end{bmatrix}$$

In general, matrix T is equal to

$$T = \begin{bmatrix} t_{L-1} & t_{L-2} & \cdots & t_0 \\ t_L & t_{L-1} & \cdots & t_1 \\ \vdots & & & \vdots \\ t_{25} & \cdots & \cdots & t_{25-L+1} \end{bmatrix}$$

and matrices C and T are of dimensions $(16+L-1 \times L)$ and $(26-L+1 \times L)$, respectively. This form of matrix T assumes perfect synchronization (i.e., exact knowledge of the length of the channel and position of the training sequence). It should be noted that matrix T will consist only of training sequence symbols as long as $26-L+1 \geq 16+L-1$, which implies that $L \leq 6$. In the case $L > 6$, matrix T must include some data samples in addition to training sequence symbols. For example,

10

$$T = \begin{bmatrix} t_5 & t_4 & \cdots & t_0 & d_{57} \\ t_6 & t_5 & \cdots & t_1 & t_0 \\ \vdots & \vdots & & \vdots & \vdots \\ t_{25} & t_{24} & \cdots & t_{20} & t_{19} \\ d_{58} & t_{25} & \cdots & t_{21} & t_{20} \end{bmatrix}$$

Thus, estimates for CIR taps 0 and 6, in this case, become:

$$\hat{h}_0 = h_0 + \frac{1}{16} (t_5 d_{57} + \sum_{j=6}^{20} t_j t_{j-6}) h_6 + f(z)$$

$$\hat{h}_6 = h_6 + \frac{1}{16} (t_{20} d_{58} + \sum_{j=5}^{19} t_j t_{j+6}) h_0 + f(z).$$

- 5 It can be seen that both the first and last CIR tap estimates will be distorted. This is much more significant in the case of the last tap, h_6 , as the first tap can, in some cases, be very large.

In matrix notation, cross-correlation CIR estimates are given by

$$\hat{h} = \frac{1}{16} Cr = \frac{1}{16} C(Th + z) = h + \frac{1}{16} Cz = h + \Delta h$$

- 10 A CIR error covariance matrix is given by

$$E[\Delta h \Delta h^H] = \frac{1}{16^2} E[Czz^H C^H] = \frac{\sigma^2}{16^2} CC^H$$

where (for $L \leq 5$, and perfect synchronization) CC^H is approximately equal to $16 I$ (I is identity matrix). If $L \geq 7$, or in case of imperfect synchronization, the estimation error covariance matrix becomes a function of CIR and data in addition to the error caused by

- 15 AWGN. In general, CIR estimation error covariance is roughly equal to $\sigma^2/16$ for all taps with cross-correlation CIR estimation.

In order to start CIR estimation, the position of the training sequence needs to be identified. The first synchronization step takes place on a once-per-symbol sampled received signal. Synchronization is performed by correlating the entire training sequence with the received signal. The result is given by

5

$$a_k = \sum_{j=0}^{25} y_{k+j} t_j = \sum_{j=0}^{25} t_j \left(\sum_{i=0}^{L-1} h_i t_{j-i+k} + z_{j+k} \right) = \sum_{i=0}^{L-1} h_i \sum_{j=0}^{25} t_j t_{j-i+k} + \sum_{j=0}^{25} t_j z_{j+k}$$

$$a_k = \sum_{i=0}^{L-1} h_i \sum_{j=0}^{25} t_j t_{j-i+k} + \sum_{j=0}^{25} t_j z_{j+k} = h_k T(0) + \sum_{\substack{i=0 \\ i \neq k}}^{L-1} h_i T(k-i) + f(t, z)$$

where $T(0)$ is the auto-correlation of the training sequence at delay of 0 (equal to 26) while $T(k-i)$ are auto-correlations of the training sequence at different delays combined with cross-correlation with data symbols. In this case, $T(k-i)$ are typically significantly smaller than $T(0)$ so that a_k is roughly equal to $h_k T(0)$. Consequently, the index of the largest magnitude a_k corresponds to the position of the training sequence start, shifted by the position of the largest magnitude CIR tap with respect to the first CIR tap (tap 0).

After finding the approximate position of the largest magnitude CIR tap, the CIR span and the actual start of the training sequence (and therefore the whole burst) need to be established. The received burst can be envisioned as a superimposition of several copies of the original data – echoes -- due to different CIR taps. As the largest magnitude CIR tap is not necessarily the first tap in the CIR, the echo corresponding to this tap does not necessarily correspond to the real start of the training sequence. Thus, a solution is needed to identify

how many taps are present in the CIR and where the CIR starts (i.e. how many taps exist in front of the largest magnitude tap).

Usually, the length of the CIR is determined in advance and is supposed to be sufficient to cover the span in the worst-case scenario. In EDGE that is the HT100 channel, with about 5 7 taps. The extent of the CIR is determined as follows: after finding the location of the largest magnitude taps, rough estimates are obtained of another ten taps on both sides of the largest magnitude tap using cross-correlation CIR estimation. These ten taps are selected based on the properties of the training sequence. Since the largest magnitude tap is a part of the CIR, the remaining taps that can be correctly estimated using cross-correlation must be on 10 either side or both sides of the largest magnitude tap. Thus, these extra ten taps are estimated, with the knowledge that at least five of them are correct. These are usually the largest magnitude taps out of the estimated eleven. Thus, the next step is usually to apply the so-called maximum energy algorithm. The magnitudes of all tentatively estimated taps are computed. Then, the CIR energy is computed given that CIR should have L taps by adding 15 the magnitudes of L taps for different valid delays. The maximum energy delay is taken as the correct CIR span. If the index of the largest magnitude tap is given by `max_idx`, and the shift of the maximum energy CIR span with respect to `max_idx` is `tap_shift`, then the start of the burst is given by `max_idx-tap_shift-61`.

In Least Squares CIR estimation, a matrix representation model for the channel is given by $r=Th+z$, where

$$T = \begin{bmatrix} t_{L-1} & t_{L-2} & \cdots & t_0 \\ t_L & t_{L-1} & \cdots & t_1 \\ \vdots & & & \vdots \\ t_{25} & \cdots & \cdots & t_{25-L+1} \end{bmatrix}$$

$h=[h_0 \ h_1 \ \dots \ h_{L-1}]^T$, $r=[r_{L-1} \ r_L \ r_{L+1} \ \dots \ r_{25}]^T$, and $z=[z_{L-1} \ z_L \ z_{L+1} \ \dots \ z_{25}]^T$.

5 Then, the LS CIR estimate is given by

$$\hat{h} = (T^H T)^{-1} T^H r = (T^H T)^{-1} T^H (Th + z) = h + (T^H T)^{-1} T^H z = h + \Delta h$$

Unlike the cross-correlation CIR estimation, the LS CIR estimation is a coupled problem. In a LS CIR estimate, every estimated CIR tap depends on other estimated CIR taps. Thus, the value of the estimated taps in the LS CIR estimate depends on the time alignment of the estimator with respect to the largest magnitude tap, inasmuch as the contribution of every omitted significant CIR tap is distributed between the remaining taps in the LS CIR estimation. Consequently, while the initial synchronization step remains unchanged with respect to cross-correlation CIR estimation, the optimal CIR alignment step is significantly different.

15 To find the optimal CIR span and alignment with respect to the largest magnitude CIR tap, all CIR estimates need to be identified with particular L that include the largest magnitude CIR tap and then pick the one that is the best fit in some sense (for example, model fitting error).

Unfortunately, this procedure is very complex -- about three to six times more complex than cross-correlation CIR estimation. Furthermore, the simplest way to reduce the complexity of the CIR alignment step, by obtaining a rough LS CIR estimate consisting of 11 CIR taps, as with cross-correlation CIR estimation, results in a CIR estimation error that is far inferior to that achievable with cross-correlation CIR estimation.

Consequently, if the optimal CIR alignment procedure is too complex, the cross-correlation CIR estimate can be used as a rough CIR estimate and thus obtain the parameters required to find a better CIR estimate, CIR length L , and the alignment of the CIR estimate with respect to the largest magnitude CIR tap. The procedure for obtaining these parameters is the same as for cross-correlation CIR estimation (described above).

For LS CIR estimation, an estimation error is given by $\Delta h = (T^H T)^{-1} T^H z$, while estimation error covariance matrix is given by:

$$E[\Delta h \Delta h^H] = E[(T^H T)^{-1} T^H z z^H T (T^H T)^{-1}] = \sigma^2 (T^H T)^{-1}$$

Error covariance for individual taps is given by diagonal elements of the estimation error covariance matrix and is roughly equal to $\sigma^2/(26-L+1)$ for LS CIR estimation. As off-diagonal elements of $(T^H T)^{-1}$ are in general not equal to zero (although small), LS CIR estimation errors are correlated, but only slightly.

Parameters obtained by channel estimation, including DC offset and CIR estimates, are further sent to data receiver and are used to enable the demodulation process (e.g. prefilter

calculation, equalization algorithm selection, calculation of the path metric of trellis based equalizers, etc.)

SUMMARY OF THE INVENTION

The present invention provides for a method for reducing DC offset associated with a receiver comprising the steps of: (a) receiving the received burst $r(n)$; (b) storing the received burst samples, $r(n)$, in a memory; (c) averaging said stored burst samples, $r(n)$, and
 5 calculating an initial DC offset, A_0 , from the stored burst samples; (d) removing DC offset value from stored burst as follows: $r(n) - A_0$; (e) identifying a rough timing estimate defining a position of largest channel impulse response (CIR) tap via cross-correlating stored burst data with a training sequence; (f) performing fine CIR synchronization to identify taps to be added to said identified largest CIR tap; (g) estimating an updated DC offset, A_1 , and a CIR, \hat{h} , via
 10 a perturbed LS CIR estimation where the received burst $r(n)$ is modeled as follows:

$$r_j = \sum_{i=0}^{L-1} h_i t_{j-i} + f_j m + z_j$$

where h_i are CIR taps, t_j are known training sequence symbols, f_j is a generic function of j , m is static DC offset; and z_j is additive white Gaussian noise, and removing updated DC offset from stored burst as follows: $r(n) - A_0 - A_1$. Perturbed LS CIR estimation can be carried in one
 15 step using the model stated above, or in two steps where first step consists of joint DC offset and CIR estimation and the second step is based on the model stated above.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 illustrates a block diagram of a data receiver in accordance with the prior art.

5 Figure 2 illustrates an overview of the present invention's method for reducing DC offsets at the receiver.

Figure 3 illustrates an overview of the present invention's "perturbed" joint DC offset and CIR estimation algorithm.

10 Figure 4 illustrates an overview of a system implemented in the receiver incorporating the methods of Figures 2 and 3.

DETAILED DESCRIPTION OF THE INVENTION

Although the present invention has been shown and described with respect to several preferred embodiments thereof, various changes, omissions, and additions to the form and
15 detail thereof may be made therein without departing from the spirit and scope of the invention.

In Direct Conversion receivers DC offset is created by the self-mixing of either the received signal or the Local Oscillator (LO) signal. In case of the received signal, signal leakage causes self-coupling via the LO path. In case of the LO signal, leakage and reflection
20 from the antenna allow a portion of the LO signal to couple via the signal path, resulting in self-mixing. Both mechanisms produce two components: one at DC and the other one at 2

times the carrier frequency. While the second component is removed by LP filtering, the DC component introduces DC offset, which, in some cases, can be very large. In general, the signal to DC offset ratio (expressed in dB) can be negative. LO-produced DC offset is usually constant over the received burst, changing slowly depending on IC temperature, IC layout, and design and other environmental factors. DC offset in Direct Conversion receivers must be somehow corrected in the baseband to maintain reliable reception. While with GMSK modulation simple DC offset compensation schemes such as simple averaging or circle center estimation are sufficient, 8-PSK modulation that is used in some of EDGE modes requires a more sophisticated DC offset compensation scheme.

Figure 2 illustrates an overview of the present invention's method **200** for reducing DC offsets at the receiver. In step **202** burst samples, $r(n)$, are received and stored locally at the receiver. Next, in step **204**, the stored burst samples, $r(n)$, are averaged and an initial DC offset, A_0 , is identified. In step **206**, the DC offset value is removed from the stored burst as follows: $r(n) - A_0$. In step **208**, a rough timing estimate of the position of largest a channel impulse response (CIR) is obtained by cross-correlating received data with a training sequence. Next, in step **210**, the maximum energy algorithm for fine CIR synchronization is used to find which taps should be added to the maximum CIR tap. Further, in step **212**, an updated DC offset, A_1 , and CIR, \hat{h} , are estimated via a perturbed LS CIR estimation representation modeling received burst $r(n)$ as follows:

$$r_j = \sum_{i=0}^{L-1} h_i t_{j-i} + f_j m + z_j$$

where h_i are CIR taps, t_j are known training sequence symbols, f_j is a generic function of j , m is static DC offset; and z_j is additive white Gaussian noise. The updated DC offset is then removed from stored burst as follows: $r(n) - A_0 - A_1$.

Figure 3 provides a general overview of the present invention's "perturbed" joint DC offset and CIR estimation algorithm 300. In step 302, a joint least squares (LS) CIR DC offset (DCO) algorithm is used to obtain initial CIR and DCO estimates. It is important to note that LS is one of the possible approaches in joint CIR DC offset estimation as outlined in Zvonar's U.S. patent (6,504,884). It is used in the description as specific computational method, though it can be replaced by other methods such as iterative computation, MMSE, etc.

Next, in step 304, the calculated DCO estimate is subtracted from the received signal burst, $r(n)$. It should be noted that steps 302 and 304 are optional, depending on the required quality of the estimate. Further, in step 306, a "perturbed" joint LS CIR DCO estimation matrix is computed based upon the training sequence used. In step 308, the LS CIR DCO estimation is repeated using the above-mentioned "perturbed" estimation matrix. Next, in step 310, the newly calculated DCO estimate is subtracted from the received signal burst, $r(n)$.

Figure 4 illustrates an overview of a system 400 implemented in the receiver incorporating the methods of Figures 2 and 3. The received signal is stored in burst storage 402. Next, burst signal, $r(k)$, stored in 402 is input into block 404, which performs a joint estimation of DC offset and CIR. DCO A_0 estimated in block 404 is forwarded to block 406, which then removes the DC offset as follows: $r(k) - A_0$. Next, the residual DC offset is estimated, in block 408, via a joint estimation of DC offset and CIR using a "perturbed"

estimation matrix, and block **410** removes the calculated residual DC offset as follows: $r(k) - A_0 - A_1$. Further, the estimated channel impulse response, h_0 , is forwarded to equalizer **412**. A detailed description of the steps of Figures 2 and 3 are now presented.

5 AVERAGING

Averaging typically removes existing DC offset. Unfortunately, however, it also introduces a DC offset that is a function of the data and channel response. If the averaging is confined to the training sequence, then the data portion of this introduced DC offset is known. The drawback of limiting averaging to the training sequence only is that the DC noise term will be larger than if the averaging is conducted over the entire burst. Considering a received signal corresponding to the training sequence,

$$r_n = m + \sum_{i=0}^{L-1} h_i t_{n-i} + z_n$$

where r_n is the received signal, m is the fixed DC offset, h_i is the channel taps, t_i is the training sequence, and z_n is the AWG noise. By averaging the signal over the training sequence, \hat{m} is given by:

$$\hat{m} = \frac{1}{M} \sum_{j=0}^{M-1} y_j = m + \frac{1}{M} \sum_{i=0}^{L-1} h_i \sum_{j=0}^{M-1} t_{j-i} + \frac{1}{M} \sum_{j=0}^{M-1} z_j$$

After subtracting the DC offset estimate from r_n ,

$$\hat{r}_n = r_n - \hat{m} = \sum_{i=0}^{L-1} h_i (t_{n-i} - \frac{1}{M} \sum_{j=0}^{M-1} t_{j-i}) + z_n - \frac{1}{M} \sum_{j=0}^{M-1} z_j$$

It should be noted that if channel estimation uses the received signal samples corrected by the average burst amplitude term, the DC offset estimation always works on the same type of problem (residual DCO which is a function of data and noise) and the algorithm can be fine-tuned.

5

JOINT LS DC OFFSET AND CIR ESTIMATION

A simple modification of the CIR model allows the joint estimation of CIR and DC offset. The general framework of joint CIR and DC offset estimation has been presented in the Zvonar's U.S. patent (6,504,884). For the purposes of this discussion, specific focus is directed towards a specific case of joint estimation using LS estimation. The received signal is rewritten as

10

$$r_j = \sum_{i=0}^{L-1} h_i t_{j-i} + m + z_j$$

where m is the DC offset. In matrix form, the above equation is rewritten as $\mathbf{r} = \mathbf{T}\mathbf{h} + \mathbf{z}$, where

15

\mathbf{T} becomes

$$T = \begin{bmatrix} t_{L-1} & t_{L-2} & \cdots & t_0 & 1 \\ t_L & t_{L-1} & \cdots & t_1 & 1 \\ \vdots & & & \vdots & 1 \\ t_{25} & \cdots & \cdots & t_{25-L+1} & 1 \end{bmatrix}$$

$h=[h_0 \ h_1 \ \dots \ h_{L-1} \ m]^T$, while r and z are unchanged. The LS CIR estimator remains unchanged, except that now a different matrix T is used to estimate $L+1$ taps. After estimation, the DC offset estimate is subtracted from the received signal.

5 SYNCHRONIZATION, CIR ALIGNMENT AND CIR LENGTH

In the presence of DC Offset, performance of all steps prior to actual CIR (plus DCO) estimation deteriorates. Some of the steps, such as initial synchronization, are robust and fairly immune to DCO, while others, such as rough cross-correlation CIR estimation, deteriorate more.

10 Initial synchronization is obtained by cross-correlating the whole training sequence with a received signal buffer. The delay that produces the largest magnitude cross-correlation is taken for the index of the largest CIR tap and the start of the training sequence realization corresponding to that tap.

The optimal fine CIR synchronization procedure with LS type estimates consists of
15 repeating LS CIR DCO estimation for different CIR offsets, including the largest magnitude tap and picking the CIR+DCO realization with maximum energy. However, this procedure would require eleven repetitions of the LS CIR DCO estimation for different delays and is excessively complex. Instead, a suboptimal fine synchronization procedure can be used wherein the cross-correlation CIR estimate is used as a rough CIR estimate.

CIR ESTIMATION ERROR

It is interesting to consider the relationship between the LS CIR estimation and the joint LS CIR DCO estimation. The LS problem to be solved is given by $T^H T h = T^H r$. In case of joint LS CIR DCO estimation, $T^H T$ is given by

$$T^H T = \begin{bmatrix} \sum_{j=L-1}^{25} |t_j|^2 & \sum_{j=L-1}^{25} t_j^H t_{j-1} & \cdots & \sum_{j=L-1}^{25} t_j^H t_{j-L+1} & \sum_{j=L-1}^{25} t_j^H \\ \sum_{j=L-1}^{25} t_j t_{j-1}^H & \sum_{j=L-1}^{25} |t_{j-1}|^2 & \cdots & \sum_{j=L-1}^{25} t_{j-1}^H t_{j-L+1} & \sum_{j=L-1}^{25} t_{j-1}^H \\ \vdots & \vdots & & \vdots & \vdots \\ \sum_{j=L-1}^{25} t_j t_{j-L+1}^H & \sum_{j=L-1}^{25} t_{j-1} t_{j-L+1}^H & \cdots & \sum_{j=L-1}^{25} |t_{j-L+1}|^2 & \sum_{j=L-1}^{25} t_{j-L+1}^H \\ \sum_{j=L-1}^{25} t_j & \sum_{j=L-1}^{25} t_{j-1} & \cdots & \sum_{j=L-1}^{25} t_{j-L+1} & 26 - L + 1 \end{bmatrix}$$

while $T^H r$ is given by

$$T^H r = \begin{bmatrix} \sum_{j=L-1}^{25} t_j^H r_j \\ \sum_{j=L-1}^{25} t_{j-1}^H r_j \\ \vdots \\ \sum_{j=L-1}^{25} t_{j-L+1}^H r_j \\ \sum_{j=L-1}^{25} r_j \end{bmatrix}$$

The last row of the equation gives an expression for DCO estimation:

$$\hat{m} = \frac{1}{26 - L + 1} \left(\sum_{j=L-1}^{25} (r_j - \sum_{i=0}^{L-1} \hat{h}_i t_{j-i}) \right)$$

Parameter m can be eliminated from the above equation, thus obtaining the expression for the estimation of CIR only:

$$\hat{h} = (T^H (T - \Delta T))^{-1} T^H (r - \bar{r})$$

where T is now of reduced dimension $(L \times 26 - L + 1)$, ΔT is a $(L \times 26 - L + 1)$ matrix given by

$$\Delta T = \frac{1}{26 - L + 1} \begin{bmatrix} \sum_{j=L-1}^{25} t_j & \sum_{j=L-1}^{25} t_{j-1} & \cdots & \sum_{j=L-1}^{25} t_{j-L+1} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ \sum_{j=L-1}^{25} t_j & \sum_{j=L-1}^{25} t_{j-1} & \cdots & \sum_{j=L-1}^{25} t_{j-L+1} \end{bmatrix}$$

while

$$\bar{r} = \frac{1}{26 - L + 1} \begin{bmatrix} \sum_{j=L-1}^{25} r_j \\ \vdots \\ \vdots \\ \sum_{j=L-1}^{25} r_j \end{bmatrix}$$

is a $(L \times 1)$ vector.

Note that it can be shown that the following relations are valid:

$$\begin{aligned} T^H \Delta T &= \Delta T^H T = \Delta T^H \Delta T \\ \Delta T^H r &= \Delta T^H \bar{r} \end{aligned}$$

Consequently,

$$\hat{h} = (T^H (T - \Delta T))^{-1} T^H (r - \bar{r}) = ((T - \Delta T)^H (T - \Delta T))^{-1} (T - \Delta T)^H (r - \bar{r})$$

and, since $\bar{r} = mU + \Delta T h + \bar{z}U$ (where $\bar{z} = \frac{1}{26-L+1} \sum z_j$ and U is the column vector with elements equal to 1), the CIR estimate becomes

$$\hat{h} = ((T - \Delta T)^H (T - \Delta T))^{-1} (T - \Delta T)^H (r - \bar{r}) = (\tilde{T}^H \tilde{T})^{-1} \tilde{T}^H (r - \bar{r})$$

while

$$r - \bar{r} = mU + Th + z - mU - \Delta Th - \bar{z}U = (T - \Delta T)h + z - \bar{z}U = \tilde{T}h + z - \bar{z}U.$$

Hence,

$$\hat{h} = (\tilde{T}^H \tilde{T})^{-1} \tilde{T}^H (r - \bar{r}) = (\tilde{T}^H \tilde{T})^{-1} \tilde{T}^H (\tilde{T}h + z - \bar{z}U) = h + (\tilde{T}^H \tilde{T})^{-1} \tilde{T}^H (z - \bar{z}U) = h + \Delta h$$

The joint LS CIR DCO estimation error covariance matrix for CIR taps therefore becomes

$$E[\Delta h \Delta h^H] = (\tilde{T}^H \tilde{T})^{-1} \tilde{T}^H E[(z - \bar{z}U)(z - \bar{z}U)^H] \tilde{T} (\tilde{T}^H \tilde{T})^{-1} = \sigma^2 (\tilde{T}^H \tilde{T})^{-1}$$

In other words, the effect of a joint LS CIR DCO estimation on CIR estimation error is that the CIR estimation error covariance matrix, $(T^H T)^{-1}$, is perturbed by the matrix $T^H \Delta T$. If the

elements of matrix $T^H \Delta T$ were small, the CIR estimation error would not be strongly affected.

Unfortunately, in most EDGE training sequences (sequences number 2 to 7) the elements of matrix $T^H \Delta T$ are quite large and contribute to the significant worsening of CIR estimation error. For example, for $L=7$ and training sequence TSC=5, the CIR estimation error is reduced from roughly $\sigma^2/20$ ($20=26-7+1$) to roughly $\sigma^2/13$, which is even worse than what is

ideally expected from the cross-correlation CIR estimation error.

As shown earlier, the DCO estimate in the joint LS CIR DCO estimation is given by

$$\hat{m} = \frac{1}{26 - L + 1} \left(\sum_{j=L-1}^{25} (r_j - \sum_{i=0}^{L-1} \hat{h}_i t_{j-i}) \right).$$

Since

$$r_j = m + \sum_{i=0}^{L-1} h_i t_{j-i} + z_j$$

the DCO estimation error is equal to

$$\Delta m = \hat{m} - m = \frac{1}{26 - L + 1} \sum_{j=L-1}^{25} \sum_{i=0}^{L-1} \Delta h_i t_{j-i} + \frac{1}{26 - L + 1} \sum_{j=L-1}^{25} z_j$$

or in matrix notation,

$$\Delta m U = \Delta T \Delta h + \bar{z} U .$$

5 The DCO tap estimation error covariance is given by

$$E[\Delta m \Delta m^H] = \text{trace}(\Delta T E[\Delta h \Delta h^H] \Delta T^H) \frac{1}{26 - L + 1} + \frac{\sigma^2}{26 - L + 1}$$

where $\text{trace}()$ denotes the sum of diagonal elements of a matrix. It is seen that the DCO tap estimation error variance is proportional to the sum of CIR taps estimation error variances and must therefore be several times larger. Indeed, with $L=7$ and training sequence TSC=5, while CIR tap estimation error variance is roughly $\sigma^2/13$, DCO tap estimation variance is roughly $\sigma^2/4.5$.

PERTURBED JOINT LS CIR DCO ESTIMATION

As shown previously, the reason joint LS CIR DCO estimation produces significantly worse performance with some training sequences is that most of EDGE training sequences were not created with joint LS CIR DCO estimation in mind. While all of them have good auto- and cross- correlation properties, means of different portions of most training sequences do not tend to zero and are actually quite large. It should be clear from the previous discussion that the performance of the joint LS CIR DCO estimation would approach that of the LS CIR estimation provided elements of matrix ΔT could be made to approach zero.

Consider, for example, that instead of the model $r_j = \sum_{i=0}^{L-1} h_i t_{j-i} + m + z_j$, the model

$$r_j = \sum_{i=0}^{L-1} h_i t_{j-i} + f_j m + z_j \text{ is used, where } f_j \text{ is a generic function of } j, \text{ for example } f_j = \sum_p e^{\frac{i2\pi j}{k_p}}$$

where p is the number of factors for the function and k_p is an integer selected by designer. It

should be noted that both the joint LS CIR DCO estimation and LS CIR estimation are special

cases of this model. In the former, $f_j=1$, while in the latter, $f_j=0$.

With the new model, our estimation problem becomes $T(f)^H T(f)h = T(f)^H r$, while $T(f)^H T(f)$ is given by:

$$T(f)^H T(f) = \begin{bmatrix} \sum_{j=L-1}^{25} |t_j|^2 & \sum_{j=L-1}^{25} t_j^H t_{j-1} & \cdots & \sum_{j=L-1}^{25} t_j^H t_{j-L+1} & \sum_{j=L-1}^{25} f_j t_j^H \\ \sum_{j=L-1}^{25} t_j t_{j-1}^H & \sum_{j=L-1}^{25} |t_{j-1}|^2 & \cdots & \sum_{j=L-1}^{25} t_{j-1}^H t_{j-L+1} & \sum_{j=L-1}^{25} f_{j-1} t_{j-1}^H \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \sum_{j=L-1}^{25} t_j t_{j-L+1}^H & \sum_{j=L-1}^{25} t_{j-1} t_{j-L+1}^H & \cdots & \sum_{j=L-1}^{25} |t_{j-L+1}|^2 & \sum_{j=L-1}^{25} f_j t_{j-L+1}^H \\ \sum_{j=L-1}^{25} t_j f_j^H & \sum_{j=L-1}^{25} t_{j-1} f_{j-1}^H & \cdots & \sum_{j=L-1}^{25} t_{j-L+1} f_{j-L+1}^H & \sum_{j=L-1}^{25} |f_j|^2 \end{bmatrix}$$

and $T(f)^H r$ is given by

$$T(f)^H r = \begin{bmatrix} \sum_{j=L-1}^{25} t_j^H r_j \\ \sum_{j=L-1}^{25} t_{j-1}^H r_j \\ \vdots \\ \sum_{j=L-1}^{25} t_{j-L+1}^H r_j \\ \sum_{j=L-1}^{25} f_j^H r_j \end{bmatrix}$$

The last row of the equation gives a modified expression for DCO estimation:

$$\hat{m} = \frac{1}{\sum_{j=L-1}^{25} |f_j|^2} \left(\sum_{j=L-1}^{25} f_j^H (r_j - \sum_{i=0}^{L-1} \hat{h}_i t_{j-i}) \right)$$

As above, parameter m can be eliminated from the above equation, thus obtaining the same expression for estimation of the CIR only:

$$\hat{h} = (T^H (T - \Delta T(f)))^{-1} T^H (r - \bar{r}(f))$$

where T is now of reduced dimension $(L \times 26 - L + 1)$, unchanged from before since it is not a function of f_j , while $\Delta T(f)$ is a $(L \times 26 - L + 1)$ matrix, given by

$$\Delta T(f) = \frac{1}{\sum_{j=L-1}^{25} |f_j|^2} \begin{bmatrix} f_{L-1} \sum_{j=L-1}^{25} f_j^H t_j & f_{L-1} \sum_{j=L-1}^{25} f_j^H t_{j-1} & \cdots & f_{L-1} \sum_{j=L-1}^{25} f_j^H t_{j-L+1} \\ \vdots & \vdots & & \vdots \\ f_{25} \sum_{j=L-1}^{25} f_j^H t_j & f_{25} \sum_{j=L-1}^{25} f_j^H t_{j-1} & \cdots & f_{25} \sum_{j=L-1}^{25} f_j^H t_{j-L+1} \end{bmatrix}$$

and

$$\bar{r}(f) = \frac{1}{\sum_{j=L-1}^{25} |f_j|^2} \begin{bmatrix} f_{L-1} \sum_{j=L-1}^{25} f_j^H r_j \\ \vdots \\ f_{25} \sum_{j=L-1}^{25} f_j^H r_j \end{bmatrix}$$

is a $(L \times 1)$ vector.

Note that, as for $f_j=1$, the following relations are still valid:

$$T^H \Delta T(f) = \Delta T(f)^H T = \Delta T(f)^H \Delta T(f)$$

$$\Delta T(f)^H r = \Delta T(f)^H \bar{r}(f)$$

Consequently,

$$\hat{h} = (T^H (T - \Delta T(f)))^{-1} T^H (r - \bar{r}(f)) = ((T - \Delta T(f))^H (T - \Delta T(f)))^{-1} (T - \Delta T(f))^H (r - \bar{r}(f))$$

5 and, since $\bar{r}(f) = mF + \Delta T(f)h + \bar{z}(f)$, where

$$\bar{z}(f) = \frac{1}{\sum_{j=L-1}^{25} |f_j|^2} \sum_{j=L-1}^{25} f_j^H z_j \begin{bmatrix} f_{L-1} \\ f_L \\ \vdots \\ f_{25} \end{bmatrix}$$

and

$$F = \frac{\sum_{j=L-1}^{25} f_j^H}{\sum_{j=L-1}^{25} |f_j|^2} \begin{bmatrix} f_{L-1} \\ f_L \\ \vdots \\ f_{25} \end{bmatrix},$$

the CIR estimate becomes

$$10 \quad \hat{h} = ((T - \Delta T(f))^H (T - \Delta T(f)))^{-1} (T - \Delta T(f))^H (r - \bar{r}(f)) = (\tilde{T}(f)^H \tilde{T}(f))^{-1} \tilde{T}(f)^H (r - \bar{r}(f))$$

while

$$r - \bar{r}(f) = mU + Th + z - mF - \Delta T(f)h - \bar{z}(f) = (T - \Delta T(f))h + m(U - F) + z - \bar{z}(f)$$

Hence,

$$\hat{h} = (\tilde{T}(f)^H \tilde{T}(f))^{-1} \tilde{T}(f)^H (r - \bar{r}(f)) = (\tilde{T}(f)^H \tilde{T}(f))^{-1} \tilde{T}(f)^H (\tilde{T}(f)h + m(U - F) + z - \bar{z}(f))$$

$$\hat{h} = h + (\tilde{T}(f)^H \tilde{T}(f))^{-1} \tilde{T}(f)^H (z - \bar{z}(f) + m(U - F)) = h + \Delta h$$

The CIR estimation error consists of two parts, one of which is due to noise while the other is due to model mismatch (i.e., assumption that f_j is not equal to 1). Joint LS CIR DCO estimation error covariance matrix for CIR taps therefore becomes

$$E[\Delta h \Delta h^H] = \sigma^2 (\tilde{T}(f)^H \tilde{T}(f))^{-1} + |m|^2 (\tilde{T}(f)^H \tilde{T}(f))^{-1} \tilde{T}(f)^H (U - F)(U - F)^H \tilde{T}(f) (\tilde{T}(f)^H \tilde{T}(f))^{-1}$$

$$E[\Delta h \Delta h^H] = \sigma^2 (T^H (T - \Delta T(f)))^{-1} + |m|^2 (T^H (T - \Delta T(f)))^{-1} T^H (U - F)(U - F)^H T (T^H (T - \Delta T(f)))^{-1}$$

The CIR estimation error covariance matrix consists of two parts, one stochastic due to AWGN, and the other deterministic, due to DCO.

Furthermore, DCO estimate is obtained by

$$\hat{m} = \frac{1}{\sum_{j=L-1}^{25} |f_j|^2} \left(\sum_{j=L-1}^{25} f_j^H (r_j - \sum_{i=0}^{L-1} \hat{h}_i t_{j-i}) \right)$$

Since

$$r_j = m + \sum_{i=0}^{L-1} h_i t_{j-i} + z_j$$

the DCO estimation error is equal to

$$\Delta m = \hat{m} - m = m \left(\frac{\sum_{j=L-1}^{25} f_j^H}{\sum_{j=L-1}^{25} |f_j|^2} - 1 \right) + \frac{1}{\sum_{j=L-1}^{25} |f_j|^2} \sum_{i=0}^{L-1} \Delta h_i \sum_{j=L-1}^{25} f_j^H t_{j-i} + \frac{1}{\sum_{j=L-1}^{25} |f_j|^2} \sum_{j=L-1}^{25} f_j^H z_j$$

or, in matrix notation,

$$\Delta m = m(f^H U / (f^H f) - 1) + U^H \frac{\Delta T(f) \Delta h + \bar{z}(f)}{U^H f}$$

where U is a vector with all elements equal to 1, and $f = [f_{L-1} \ f_L \ \dots \ f_{25}]^T$.

As in the case of CIR estimation error, the DCO tap estimation error also includes two parts, one due to AWGN, and the other due to model mismatch and proportional to the DCO value.

The expressions for CIR estimation error and DCO estimation error provide guidance with regard to which conditions function f_j must fulfill to reduce the estimation error.

The first condition, leading to reduction of CIR estimation error due to AWGN, is that the elements of $\Delta T(f)$ should be close to zero (or equal to zero):

$$\sum_{j=L-1}^{25} f_j^H t_{j-k} \rightarrow 0, \forall k = (0, 1, \dots, L-1)$$

If that condition is valid, the CIR estimation error covariance matrix becomes:

$$E[\Delta h \Delta h^H] = \sigma^2 (T^H T)^{-1} + |m|^2 (T^H T)^{-1} T^H (U - F)(U - F)^H T (T^H T)^{-1}$$

Estimation error due to AWGN is now the same as in the LS CIR estimation. However, the error due to model mismatch remains and is proportional to

$$26 - L + 1 - \frac{\left| \sum_{j=L-1}^{25} f_j \right|^2}{\sum_{j=L-1}^{25} |f_j|^2}$$

so that the second condition is to make the above relation also approach zero. If function f_j fulfills both of the above conditions, then the CIR estimation error for joint LS CIR DCO estimation would be the same as the CIR estimation error for LS CIR estimation.

Additionally, f_j must fulfill the following conditions to reduce the DCO estimation error.

$$\frac{\sum_{j=L-1}^{25} f_j^H}{\sum_{j=L-1}^{25} |f_j|^2} - 1 \rightarrow 0$$

This condition would eliminate error due to model mismatch. Furthermore, if $\Delta T(f)=0$, the coupling of the CIR estimation error to the DCO estimation error would also be eliminated, so the DCO estimation error would become a function of AWGN only.

$$\Delta m = \frac{1}{\sum_{j=L-1}^{25} |f_j|^2} \sum_{j=L-1}^{25} f_j^H z_j$$

In that case, DCO estimation error variance would be equal to:

$$E[\Delta m \Delta m^H] = \sigma^2 \frac{1}{\sum_{j=L-1}^{25} |f_j|^2}$$

which is the last condition for optimality and should ideally be equal to $\sigma^2/(26-L+1)$, which is the Cramer-Rao bound for LS CIR DCO estimation.

Thus, provided a function f_j can be identified that can fulfill all of the above conditions, the optimal estimation error can be found regardless of the training sequence.

Unfortunately, some of listed conditions are contradictory (i.e., cannot be simultaneously fulfilled), indicating that Cramer-Rao bound cannot be reached with perturbed joint LS CIR DCO estimation.

The paper of Hui entitled, “Using a direct conversion receiver in EDGE terminals-a new DC offset compensation algorithm,” proposes a different approach in the elimination of error caused by DC offset. The proposed method modifies the receiver in such a way that the received signal can be modeled as

$$r_j = me^{i\delta} + \sum_{k=0}^{L-1} h_k t_{j-k} + z_j$$

where $i=(-1)^{0.5}$ and δ is a pre-selected angle, a function of the training sequence. Thus, their proposal can be interpreted as a special case using function $f_j=e^{i\theta_j}$. This function is well suited for minimizing $\Delta T(f)$ and can therefore provide estimation error close to LS CIR estimation error. Hui proposes that the DCO be “rotated” in the following manner: frequency of LO which is part of radio section is to be deliberately offset by δ before the DC offset is introduced in the signal chain. When this rotation is corrected in the baseband, the DCO is effectively rotated by the same angle, since it was added to the received signal in the meantime. The improvement of the estimation error will depend on the accuracy with which the LO frequency can be controlled and the receiver requires a control path between baseband receiver (where the received data are decoded and the receiver discovers which training sequence it is supposed to receive) and the LO. Also, proposed correction will require that there is known relationship between the number of TS and introduced frequency offset, i.e. function f depends on the training sequence. Hui’s method requires interactions between radio and baseband portions of receiver, and has training sequence dependant frequency offset.

Conditions needed for the reduction of estimation error as derived by the inventors to the Cramer-Rao bound cannot all be fulfilled at the same time and problem can be approached with several suboptimal techniques.

In general, suboptimal approaches would try to find functions that fulfill some, instead of all, the above-listed conditions for optimality and modify the estimation procedure in such a way that would maximally exploit the characteristics of the perturbed joint LS estimator. Secondly, when searching for an optimal function f_j , one is faced with the difficulty of finding a closed form solution that would fulfill all conditions for optimality. However, one may opt to find closed form solutions for functions that fulfill at least some of the conditions for optimality. Alternatively, a parametric form for the function f_j can be specified, and then a search can be performed for identifying parameters that minimize some metric related to the overall estimation error.

Another approach dispenses with the seeking of a function f_j that fully eliminates errors due to mismatch and non-zero ΔT elements. Instead, a function f_j is identified that improves CIR estimation error by reducing ΔT elements while minimally increasing estimation error due to mismatch. Good solutions are found by minimizing an error-related metric with respect to parameters of some parametric function. A good candidate for such a function is obviously one or a combination of several complex sinusoids, $e^{i\delta_j}$.

One possibility is to seek a parameter δ that minimizes the metric

$$(T^H (T - \Delta T(f)))^{-1} + w(T^H (T - \Delta T(f)))^{-1} T^H (U - F)(U - F)^H T (T^H (T - \Delta T(f)))^{-1}$$

where $f_j = 1 - e^{i\delta_j}$, and w is a weight proportional to the ratio of DCO magnitude and AWGN variance. The idea is to find the function f_j that reduces CIR estimation error. Then, given that DCO estimation error is proportional to CIR estimation error, DCO is estimated as if $f_j = 1$,

$$\hat{m} = \frac{1}{26 - L + 1} \left(\sum_{j=L-1}^{25} (r_j - \sum_{i=0}^{L-1} \hat{h}_i t_{j-i}) \right)$$

Another possible approach is based on the observation that as long as the DCO estimation error in perturbed LS solution can be reduced below the level of the DCO estimation error with $f_j = 1$, iterative joint perturbed LS CIR DCO estimation (as presented in Fig. 3) is bound to reduce the overall estimation error. It turns out that for TSC=5, function $f_j = 0.5(1 - e^{i\delta_j})$, where $\delta = -2\pi/(26-L+2)$, is a good choice for such a function. Interestingly, this function fulfils the condition that

$$\frac{\sum_{j=L-1}^{25} f_j^H}{\sum_{j=L-1}^{25} |f_j|^2} - 1 \rightarrow 0$$

while also significantly reducing the elements of ΔT matrix. With this approach, an “ordinary” joint LS CIR DCO estimate is first obtained, then, as a second step, the DCO estimate is subtracted. Then, the estimation is repeated, but this time using perturbed joint LS CIR DCO estimation. The final DCO estimate is the sum of the initial and improved DCO estimates, while the CIR estimate is the second, perturbed LS CIR estimate. This method

works best at high E_b/N_0 , giving an improvement of 2.0dB in a MCS7 HT100 channel, with TSC=5, for example.

Furthermore, the present invention includes a computer program code based product, which is a storage medium having program code stored therein which can be used to instruct a computer to perform any of the methods associated with the present invention. The computer storage medium includes any of, but not limited to, the following: CD-ROM, DVD, magnetic tape, optical disc, hard drive, floppy disk, ferroelectric memory, flash memory, ferromagnetic memory, optical storage, charge coupled devices, magnetic or optical cards, smart cards, EEPROM, EPROM, RAM, ROM, DRAM, SRAM, SDRAM, and/or any other appropriate static or dynamic memory or data storage devices.

Implemented in computer program code based products are software modules for: (a) receiving a signal burst $r(n)$; (b) storing received burst samples, $r(n)$, in memory; (c) averaging the stored burst samples, $r(n)$, and calculating an initial DC offset, A_0 , from the stored burst samples; (d) removing DC offset value from stored burst as follows: $r(n) - A_0$; (e) estimating an updated DC offset, A_1 , and a channel impulse response (CIR), \hat{h} , via a perturbed LS CIR estimation representation modeling received burst $r(n)$ as follows:

$$r_j = \sum_{i=0}^{L-1} h_i t_{j-i} + f_j m + z_j$$

where h_i are CIR taps, t_j are known training sequence symbols, f_j is a generic function of j , m is static DC offset; and z_j is additive white Gaussian noise, and removing updated DC offset from stored burst as follows: $r(n) - A_0 - A_1$.

What is claimed is: